

## I. Introduction

I arrived at the Bell Telephone Laboratories in September, 1961, with the following musical and intellectual baggage:

1. numerous instrumental compositions reflecting the influence of Webern and Varèse;
2. two tape-pieces, produced in the Electronic Music Laboratory at the University of Illinois — both employing familiar, “concrete” sounds, modified in various ways;
3. a long paper (“Meta+Hodos, A Phenomenology of 20th Century Music and an Approach to the Study of Form,” June, 1961), in which a descriptive terminology and certain structural principles were developed, borrowing heavily from Gestalt psychology. The central point of the paper involves the clang, or primary aural Gestalt, and basic laws of perceptual organization of clangs, clang-elements, and sequences (a higher order Gestalt unit consisting of several clangs);
4. a dissatisfaction with all purely synthetic electronic music that I had heard up to that time, particularly with respect to timbre;

5. ideas stemming from my studies of acoustics, electronics and — especially — information theory, begun in Lejaren Hiller's classes at the University of Illinois; and finally
6. a growing interest in the work and ideas of John Cage.

I leave in March, 1964, with:

1. six tape compositions of computer-generated sounds — of which all but the first were also composed by means of the computer, and several instrumental pieces whose composition involved the computer in one way or another;
2. a far better understanding of the physical basis of timbre, and a sense of having achieved a significant extension of the range of timbres possible by synthetic means;
3. a curious history of renunciations of one after another of the traditional attitudes about music, due primarily to a gradually more thorough assimilation of the insights of John Cage.

In my two-and-a-half years here I have begun many more compositions than I have completed, asked more questions than I could find answers for, and perhaps failed more often than I have succeeded. But I think it could not have been much different. The medium is new and requires new ways of thinking and feeling. Two years are hardly enough to have

become thoroughly acclimated to it, but the process has at least been begun.

I want to express my gratitude to Max Mathews, John Pierce, Joan Miller, and to all my friends and co-workers who have done so much to make my stay here not only instructive but pleasant. My questions and requests for assistance have always been responded to with great generosity, and I shall not soon forget this.

## II. The *Noise Study*, November-December, 1961

My first composition using computer-generated sounds was the piece called *Analog #1: Noise Study*, completed in December, 1961. The idea for the *Noise Study* developed in the following way: For several months I had been driving to New York City in the evening, returning to the Labs the next morning by way of the heavily traveled Route 22 and the Holland Tunnel. This circuit was made as often as three times every week, and the drive was always an exhausting, nerve-wracking experience, fast, furious, and “noisy.” The sounds of the traffic — especially in the tunnel — were usually so loud and continuous that, for example, it was impossible to maintain a conversation with a companion. It is an experience that is familiar to many people, of course. But then something else happened, which is perhaps not so familiar to others. One day I found myself listening to these sounds, instead of trying to ignore them as usual. The activity of listening, attentively, to “non-musical,” environmental sounds was not new to me — my esthetic attitude for several years had been that these were potential musical material — but in this particular context I had not yet done this. When I did, finally, begin to listen, the sounds of the traffic became so interesting that the trip was no longer a thing to be dreaded and gotten through as quickly as possible. From then on, I actually looked forward to it as a source of new perceptual insights. Gradually, I learned to hear these sounds more acutely, to follow the evolution of single elements within the total sonorous “mass,” to feel, kinesthetically, the characteristic rhythmic articulations of the various elements in combination, etc. Then I began to try to analyze the sounds, aurally, to estimate what their physical properties might be — drawing

upon what I already knew of acoustics and the correlation of the physical and the subjective attributes of sound.

From this image, then, of traffic noises — and especially those heard in the tunnel, where the overall sonority is richer, denser, and the changes are mostly very gradual — I began to conceive a musical composition that not only used sound elements similar to these, but manifested similarly gradual changes in sonority. I thought also of the sound of the ocean surf — in many ways like tunnel traffic sounds — and some of the qualities of this did ultimately manifest themselves in the *Noise Study*. I did not want the quasi-periodic nature of the sea sounds in the piece however, and this was carefully avoided in the composition process. Instead, I wanted the a-periodic, “asymmetrical” kind of rhythmic flow that was characteristic of the traffic sounds.

The actual realization of this image in the *Noise Study* took place in three stages: first, an “instrument” was designed that would generate bands of noise, with appropriate controls over the parameters whose evolution seemed the most essential to the sonorities I had heard; second, the large-scale form of the piece was sketched out, in terms of changing mean-values and ranges of each of the variable parameters; third, the details — the actual note-values in each parameter — were determined by various methods of random number selection, “scaled” and/or normalized in such a way that the note-values fell within the “areas” outlined in step 2; fourth, these note-values, in numerical form, were used as the input “score” for the music program, containing the “instruments” designed in the first step, and a digital tape was generated and converted into analog

form; fifth, this tape was mixed with the same tape re-recorded at one-half and double speeds, for reasons — and in a way — that will be described below.

1) The Instrument (see Figure 1): The instrument is designed to produce noise-bands by random amplitude-modulation of a sinusoidal carrier, with provisions for continuous, linear interpolation between an initial and a final value (for each “note”) in amplitude, bandwidth and center frequency. (The possibility of varying the form of the carrier wave was not used in the *Noise Study*, because it was found that the sounds resulting from modulation of other wave-forms (richer in harmonics) had a peculiar quality — more like radio “static” than the sounds I was after). In addition, for the generator controlling the amplitude envelope (U1), functions other than the linear interpolation function could be specified (in which case the C4 input to U2 was set to zero). In the second half of the tape, two such functions are used, shown in Figure 2. Five of these instruments were used in the “orchestra” for this piece — all of them sounding simultaneously (though rhythmically independent) on each tape. Thus, after the three versions of the tape (at three speeds) had finally been combined, the density of independently varying noise-bands was as high as fifteen. Because of the diffuse quality of most of the sounds, it is not possible (nor was it expected) that each of these fifteen “voices” could be heard separately. The high density is nevertheless essential to the total sonority, which would (and does) sound perceptibly different with fewer voices sounding (this is one of the reasons why I mixed the three tapes in the final version).

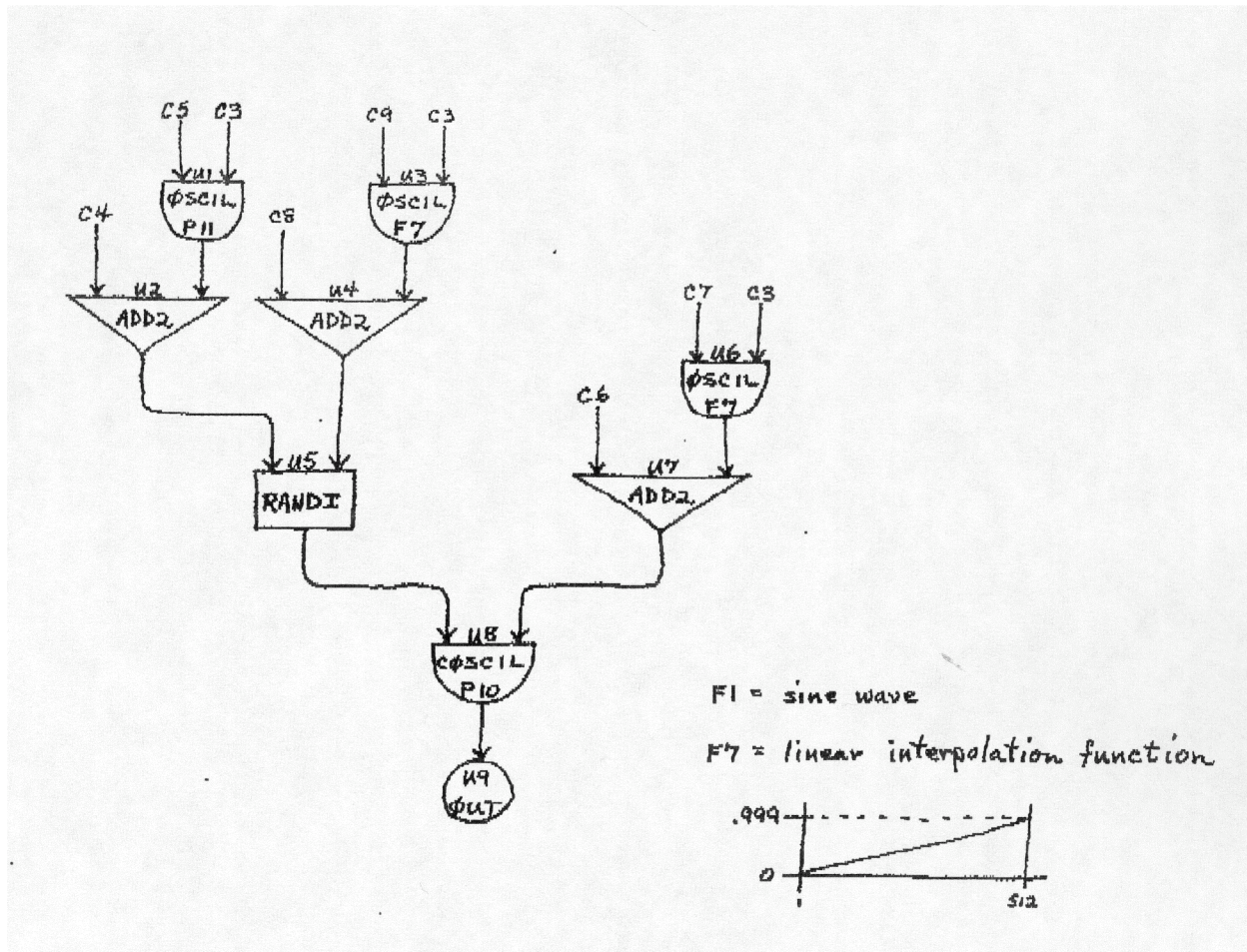


Figure 1. Instrument for *Analog #1 — Noise Study*.

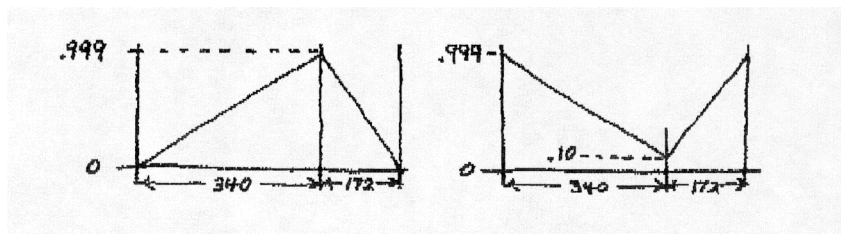


Figure 2. Amplitude-envelope functions for the *Noise Study*.

2) The Formal Outline (see Figure 3): The piece is divided into five sections, the durations of the sections decreasing, progressively, from the first to the fifth. The piece begins slowly, softly, with relatively wide noise-bands whose center frequencies are distributed evenly throughout the pitch range, approximating a white noise. As the average intensity

and temporal density increase (in the second and third sections) the noise bandwidths decrease, until the sounds of each instrument are heard as tones with amplitude fluctuations, rather than as noise-bands. The beginning of section 4 is marked by a sudden change to a lower temporal density (i.e., longer note-durations), wider bandwidths, and a new amplitude envelope is introduced, with percussive attack followed by a decreasing - then increasing - amplitude. During this fourth section the average intensity is maintained at a high level. The fifth section begins at a lower intensity, which decreases steadily to the end of the piece. This return to the conditions of the beginning of the piece is manifested in the other parameters also, except for temporal density, which increases during the last two sections from a minimum (like the beginning) to a maximum at the end. Thus, except for this note-duration parameter, the overall shape of the piece is a kind of arch.



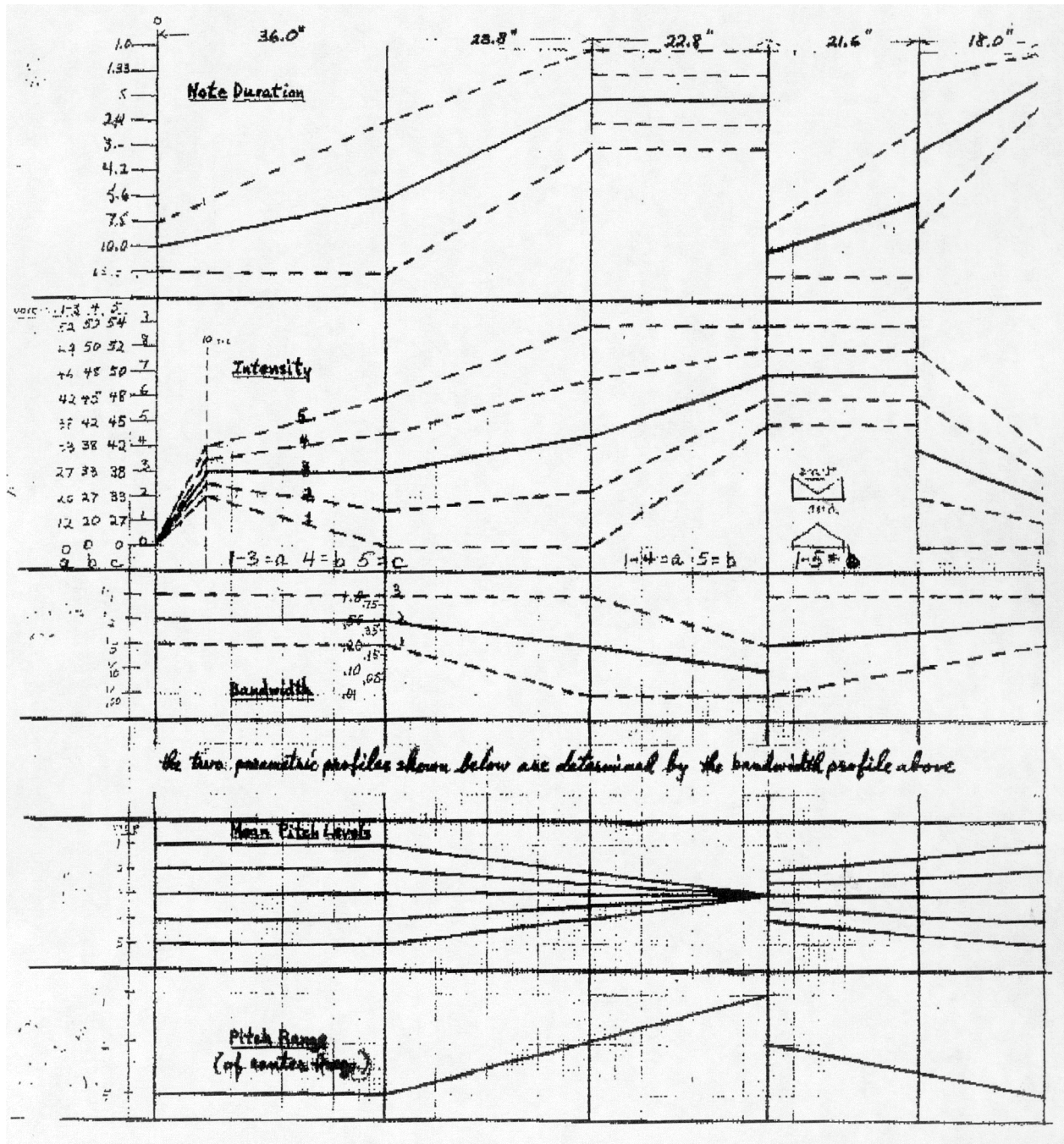


Figure 3. Parametric Means and Ranges in the *Noise Study*.

3) Determination of the "Details": Various means of random number selection were used in this stage, the method used depending on the number of quantal steps in each parametric scale and/or (what amounts to about the same thing) the number of decimal points of precision

wanted in the specifications of parametric values. For center frequency, the toss of a coin was used to determine whether the initial and final values for a given note were to be the same or different (i.e., whether the pitch of the note was constant or varying). In order to realize the means and ranges in each parameter as sketched in the formal outline, a rather tedious process of scaling and normalizing was required that followed their changes in time. A more detailed description of this does not seem of much interest here, however.

4) and 5): The fourth stage involved the standard procedures for generating the sounds specified by the “score” (as described in my article in the (Yale) *Journal of Music Theory*). The resulting analog tape seemed “successful” on first hearings, but later I began to feel somewhat dissatisfied with it in two respects: first, I would have liked it to be denser (vertically) or cover a wider range of vertical densities; and second, the range of temporal densities (speeds, note-durations) seemed too narrow – the slow sections did not seem slow enough, nor the fast sections fast enough. (I was to continue to make this mistake — especially the underestimation of the average note-durations needed to give the impression of “slowness” — for several months. Only in the most recent compositions have I finally adjusted my sense of the correlation here between the numbers representing note-duration and my subjective impression of temporal density.)

After some consideration of these problems, a very simple solution occurred to me which corrected both conditions in one stroke, though it introduced some new conditions that deviated from the original formal

outline. The original analog tape was re-recorded at half speed and at double speed, and these mixed with the original. The entrances of the three tapes were timed in such a way that the points of division between sections three and four were synchronized, thus disturbing the general shape of the piece as little as possible in the mixed version (see Figure 4, showing the temporal-density and intensity graphs of the three strata as they would appear in time). This device, while sure to antagonize certain purists, and undertaken with some hesitation on my part, seemed to give me more nearly what I was after — to correspond more closely to the original image — than the first analog tape by itself, and this is its final form. So far, no one listening to the piece has even noticed the repetitions (at different speeds and in different octaves) that resulted from the overlay — though they are plain to my ear, and will surely be heard by anyone told about it in advance.

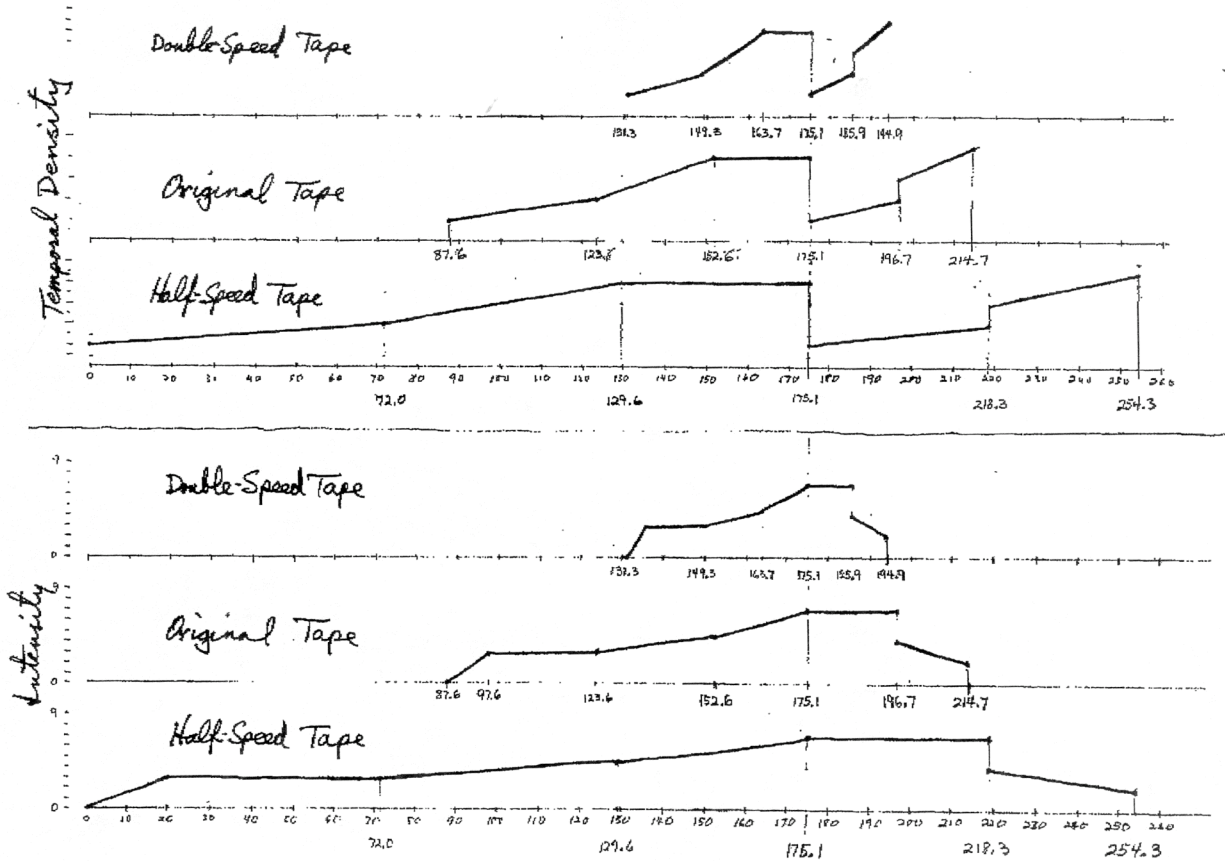


Figure 4. Superimposition of the three analog tapes in the *Noise Study*.

When the *Noise Study* was put on the “Music from Mathematics” record, the recording engineers put it through the artificial reverberation process that is used (with such bad effect, usually) on most commercial recordings. Here, to my surprise, the added reverberation had a very good effect, so I intend one day to add reverberation to the original tape itself.

### III. Psychoacoustic Experiments

Between the *Noise Study* and the *Stochastic Studies* described in Section IV, there was a period of more than a year during which no compositions were completed; a number of pieces were begun or planned, but all were abandoned before they were finished. Most of the time was spent in experiments and tests of various kinds, which will be described here under two headings: 1) modulation, and 2) rise-time.

1) Modulation: Early tests served very quickly to establish approximate limits of the rate and range of a periodic frequency modulation corresponding to the vibrato in conventional musical instruments and the voice. I found that, with sinusoidal modulation of a simple tone in the mid-range of the frequency scale, ranges of from about  $\pm.25\%$  to  $\pm 2.0\%$  (times the center frequency), at rates of 6.5 to 9.0 cycles per second were usable, with mean (or “modal”) values for these parameters at about  $\pm 1.0\%$  at 7.5 to 8.0 cps. These define the “limits” for the vibrato in this sense: a deviation from the center frequency of less than .25% is hardly perceived at all; while one greater than 2% sounds “rough” (at the fastest vibrato rates) or “wobbly” (at slower rates); at a rate slower than 6.5/sec., the successive vibrato swings are heard as “changes in frequency” as such, rather than “fusing” together into a homogeneous sound (Seashore’s “sonance”), while at rates higher than 9.0/sec., the sound is (again) “rough,” if the range is wide enough to be perceived at all.

The “optimum” values for range and rate of the vibrato seem to be somewhat different for different people, however, “good vibratos” used by others here at the Labs usually sound either too slow or too wide to my ears, and a comparison of my results with Seashore’s measure of average rates and ranges of vibratos in tones of singers shows the same disparity. That is, his singers’ vibratos are nearly all either slower or wider (or both) than a vibrato that would sound best to me with the synthetic tones. In this case, the disparities may be due simply to differences of taste (I haven’t heard the tones he measured, so I don’t know whether they would actually sound poor to me), but it might also be due to differences in other attributes of the tones (the singers’ tones were richer in harmonics and had more or less constant formant frequencies, while the synthetic tones I had been working with were usually simpler, and their spectra were modulated as a whole, “in parallel,” any formant peaks changing along with the fundamental).

The tones produced with such a periodic frequency modulation were still not very interesting, however (and the reason for studying modulation in the first place was precisely to enrich the quality of the tone, in a way suggested by conventional musical sounds). Consideration of the way “natural” tones were shaped (by a singer, for example) led to redesigning the test instruments in such a way that the vibrato parameters themselves could be made to vary in time, during the course of the tone, instead of remaining constant. Of the various possible ways of doing this, the one that seemed to correspond most closely to a conventionally “good musical tone” was the result of enveloping the vibrato range, so that it built up to its maximum toward the middle of the tone, and then decreased again

toward the end, as shown in Figure 5. Corresponding envelopes on the vibrato rate did not seem to be of much interest, probably because the range of usable vibrato rates is so much narrower than that of usable (vibrato) ranges.

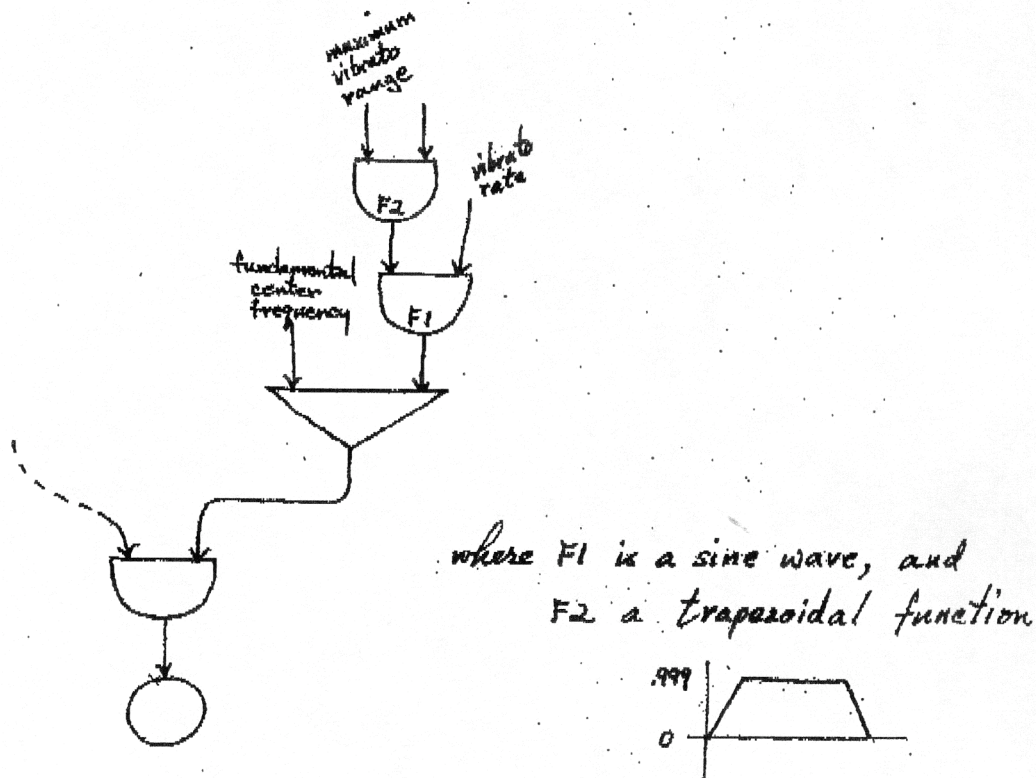


Figure 5. Enveloped vibrato-range.

A sort of “mechanical” quality still persisted in these tones, however, and in order to overcome this I began to experiment with random frequency modulation, both with and without some amount of periodic modulation. The nature of the interpolating random number generator is such that, in order to give the impression of a modulation of a range and rate similar to the periodically modulated tone, higher values in both parameters are necessary ( $\pm 5$  to 2.0%, at 16-20/sec.). Using random modulation by itself

produces an interesting tone, but it does not sound like a conventional “musical tone” with normal vibrato. The combination of random and periodic modulation, with enveloping on the ranges of each (as described above), does, however, produce an effect so “realistic” that I felt I had achieved one of the partial goals I had set for myself in these tests, when I heard the results. The relative proportion of the range allotted to the two modulation sources does not seem to make very much difference, just so long as there is a “perceptible” amount of each, and the sum of the two ranges does not exceed the range considered “good” for a periodic modulation above (about  $.5 + .5 = 1.0\%$  in my work).

With amplitude modulation, I found that the effect of a periodic modulation was not very interesting, did not even seem to be needed with the more interesting random amplitude modulation, to simulate the kind of fluctuations of amplitude that give “life” to most instrumental and vocal sounds. Only with such sounds as those of the flute, vibraphone and bell does a periodic modulation of amplitude seem perceptually important.

The useful ranges and rates of random amplitude modulation are from about  $\pm .15$  to  $\pm .50$  (times the mean amplitude), at rates of from about 4 to 30 per second. The wider ranges given reflect the greater size of the DLs for amplitude (by comparison with those for frequency), but the greater range of AM *rates* requires some explanation.

Our perception of amplitude apparently differs from the perception of frequency in such a way that the condition of “fusion” or “sonance” does



not apply here. That is, the very slow rates (4-6/sec.) are heard simply as a kind of amplitude envelope on the tone, giving it shape, not felt as a “deviation” in its primary characteristics. The faster rates (12-30/sec.) are, at the same time, quite usable for the production of “good” tone, provided that the range of AM used is small enough to avoid “roughness.” Thus, there is a kind of reciprocal relation between the range and rate of amplitude modulation that will produce a tone of ordinary “musical” character: narrow ranges with faster rates, and slower rates with wider ranges. (This reciprocal relation was later built into the PLF-3 composing program, described in Section IV.) Since the AM range is automatically enveloped in the computer-instrument, along with the main amplitude of the tone, it was not found necessary to envelope the AM range in any additional way (corresponding to that used with FM).

When random amplitude modulation is applied to the synthesized tone along with the combination of periodic and random frequency modulations already described, the result is a quality of tone that compares very favorably with that of a tone produced by a conventional musical instrument; it no longer seems “mechanical,” “lifeless,” “electronic,” etc., adding that element of richness to the computer sounds that I had so long felt necessary. Since these experiments, every instrument I designed – with the intention of producing interesting tones, employed these modulations. Figure 6 shows a typical instrument in which these modulations are all used.

The modulations effected by such an instrument (as diagrammed in Fig. 6) are applied to the signal waveform as a whole, so that all spectral

components will be modulated together, synchronously. This is an artificial condition, and I was interested to discover whether independent modulations of spectral components would enrich the tones still further. This was found to be so, but the differences were really quite small, while the generating time was considerably increased, and I have not used such independent modulations in actual compositions, primarily for this (economic) reason.

Among the various ways that spectral components may be made independent, with respect to modulation, the simplest one to work with breaks the tone up into two parts – one including odd partials only, the other even partials. The periodic frequency modulation is common to both, but the random modulations are independent. Such a tone sounds as rich as one divided into three groups of partials in various ways, so I conclude that no more than two groups are necessary. Care must be taken, though, that the range of the random frequency modulations is not too wide, because this can result in a sound like the “mistuned unison,” of two instruments playing together but only approximately in tune. (Of course, if such an effect is wanted, this is a relatively easy way to get it.)

With larger values of range and/or rate for the random generators in Fig. 6, the result will be a band of noise, with relative amplitude and bandwidth depending on the input parameters. Thus, increasing the AM rate will produce a noise-band of increasing bandwidth, centered at the tonal frequency and superimposed on the tone, as shown in Figure 7 where the relative amplitude of the noise is determined by the right-hand or “M” input to the random generator U1 in Figure 6. If A1 of U2 is set to

zero, this pure tonal component is removed, and only the noise-band remains.

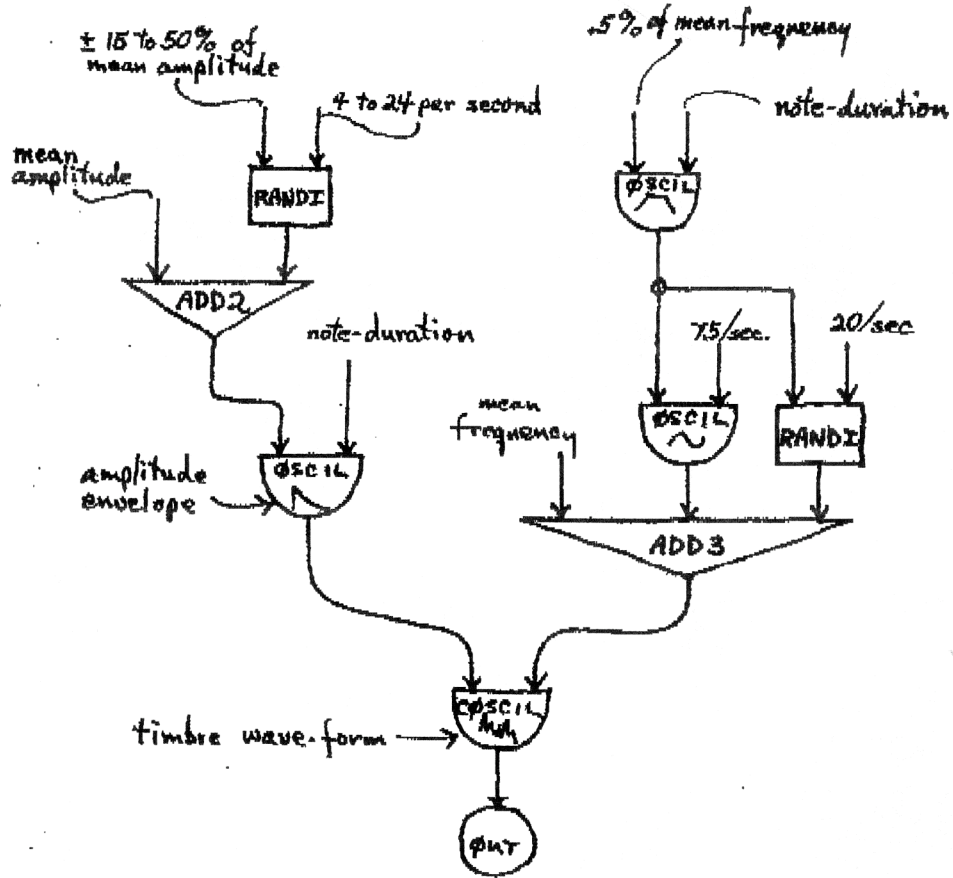


Figure 6. Instrument with slight random modulations.

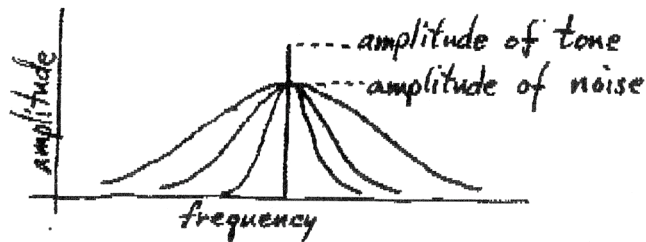


Figure 7. Noise spectra produced by random amplitude modulation.

With frequency modulations, the relations between the input settings and the characteristics of the noise-band are different, as described in my *Journal of Music Theory* article. Here, the bandwidth of the noise depends primarily on the range of the FM (rather than on the rate, as with AM), while the rate of the FM has an effect on the quality of the sound that is difficult to describe, though the differences are quite perceptible, at least among the relatively slower rates (at fast rates they are not so easily perceptible). Roughly, however, they are this: for a given random FM range(= band-width), the slower rates (30-100) result in a greater “roughness,” the sound becoming “smoother” (more homogeneous) as the rate is increased.

Acoustic analyses of both speech and singing have shown that there are irregular fluctuations of period-length (frequency) at rates as high as the mean fundamental frequency of the tone itself, though these fluctuations may cover only a very narrow range. In addition, experiments have shown that such fluctuations – in the case of speech at least – are essential to “naturalness” of the speech sounds. They contribute a kind of “noisy” character to the sounds, but the noise is of a very narrow bandwidth, and it is very probable that the timbres of many conventional musical instruments are characterized by similar, fast, narrow, quasi-random modulations. For reasons of economy, again, I have not made use of such modulations in my compositions yet, but I suspect that any attempts to simulate the sounds of conventional musical instruments would find these necessary, in addition to the slower modulations I have described (and used).

The noises that can be produced by an instrument like that drawn in Fig. 6 are centered around the frequency of the tone, as specified by A1 of U7, or around integral multiples of that frequency (harmonics). In order to generate sounds in which the noise component has a center frequency different from that of the tone, a more complex instrument-design would be necessary.

2) The Rise-time of a Tone: Instead of describing this work here, I am including, among other articles I have written here at the Labs, the paper given at a meeting of the Acoustical Society in May, 1962. The following remarks will assume a reading of that paper, or at least of the conclusions.

In retrospect, several things need to be said about the rise-time experiment. It has gradually become evident that musical context has such a powerful effect on the differential perception of rise-time and other parameters that the results of an experiment like this one are of very little use, musically. I find that in most actual musical situations, I can distinguish – at most – about three rise-times: “short,” “medium,” and “long.” Furthermore, I find the use of a scale of discrete steps in any parameter no longer necessary, and of much less interest than the use of a continuous scale, letting the ear of the listener do the “quantizing.” This the listener’s ear will do anyway, so it is a question simply of lessening the disparity between the process of composition and that of listening. One result of the experiment is useful, however: the implication of an approximately logarithmic (rather than linear) spacing on the continuum of perceived rise-times. Nearly all the parametric continua relevant to

sounds show this logarithmic condition, and my later composing programs have treated them in this way.

It is questionable whether such tests as the one described, carried out in very artificial laboratory conditions, and divorced from any musical context, can ever be of much use to the composer. And for this reason, primarily, I have not done any more experiments of this kind. Instead, I have tried to gain an understanding of such physical-to-psychological correlations more directly - by listening to the sounds in a musical context. What this approach lacks in precision (and - sometimes, unfortunately - communicability), it more than makes up for in efficiency. Only after giving up all intentions of dealing with these problems in the strict ways of the psycho-physical laboratory has it been possible for me to produce compositions with any degree of fluency.

#### IV. Four Stochastic Studies and Dialogue

If I had to name a single attribute of music that has been more essential to my esthetic than any other, it would be variety. It was to achieve greater variety that I began to use random selection procedures in the *Noise Study* (more than from any philosophical interest in indeterminacy for its own sake), and the very frequent use of random number generation in all my composing programs has been to this same end. I have tried to increase this variety at every Gestalt “level” – from that of small-scale fluctuations of amplitude and frequency in each sound (affecting timbre), to that of extended sequences of sounds – and in as many different parameters of sound as possible (and/or practicable). The concept of entropy has been extremely useful as a descriptive “measure” of variety, and several important laws of musical structure have been derived in terms of entropy relations (see the memo, “On Certain Entropy Relations in Musical Structure” included with my articles). The composing programs described below represent various attempts to combine the clang concept developed in Meta+Hodos with more recent ideas about these entropy relations and stochastic processes in general.

During the spring and summer of 1962 I designed several very elaborate instruments that generated, automatically, random sequences of tones. This was done by means of the “RANDH,” non-interpolating random number generator, modulating very long “notes.” Figure 8 shows such an instrument, in which note-duration, amplitude and frequency are all varying randomly (on linear scales, note!).

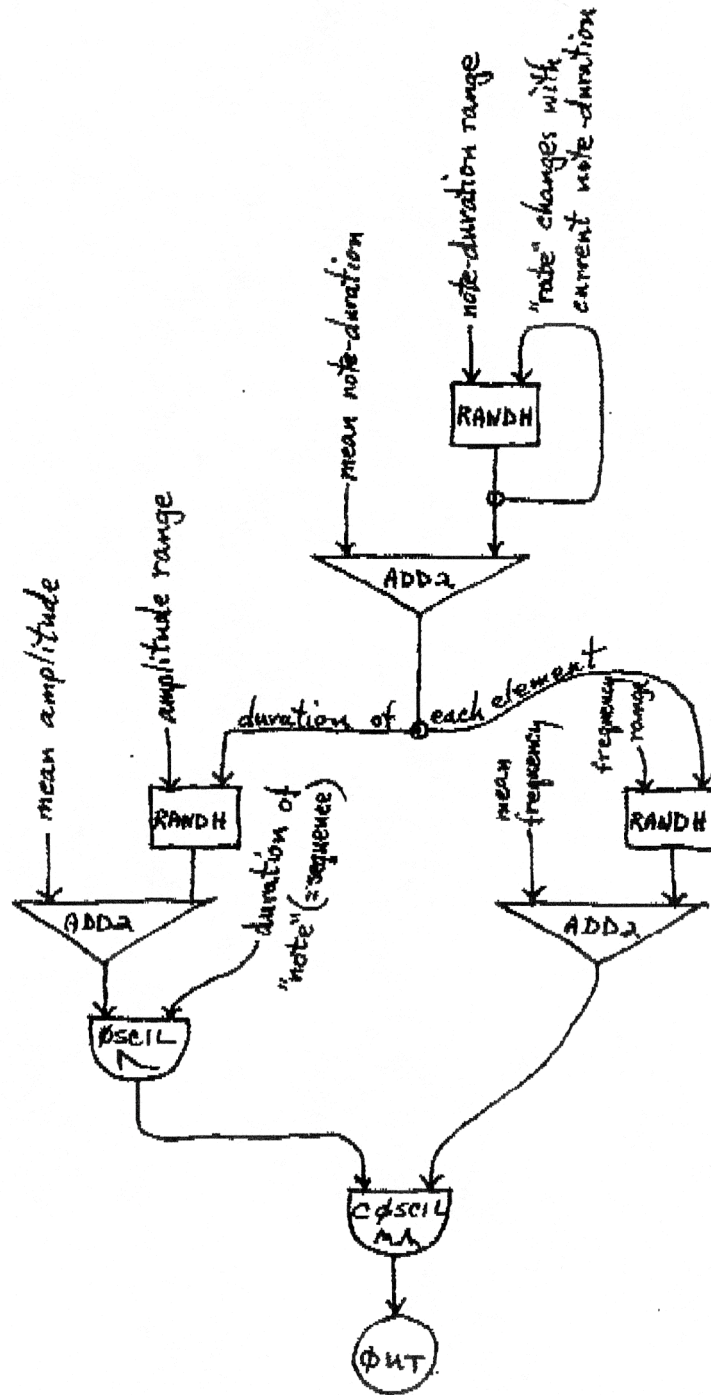


Figure 8. Instrument for generating random sequences.

Tests with these instruments produced results that were quite interesting to me, but it was not very efficient to use the compiler itself for these



operations. It became clear that programming facilities were needed that would make it possible to derive a computer “score” from another – “composing” program, maintaining a separation between the compositional procedures and the actual sample-generation. In October, 1962, Max Mathews completed the subroutines necessary for linking such composing programs to the compiler, and helped me write my first “Stochastic Music” program (“PLF 2”).

The conditions I wanted to be incorporated into this program were these: three parameters — note-duration, amplitude and frequency — were to vary randomly from note to note, but the mean-value and range of deviation around this mean was to change (also in a quasi-random way) after every second or two (i.e., from clang to clang). In addition, in each clang, at least one of the three parameters should be variable over its entire range, whereas the other parameters might be varying (temporarily) over a narrower range. No further constraints were placed on the process.

Accordingly, the input data to this program included lists of nine “states” — means and ranges (on log scales) — for each parameter, the first state listed being the one with maximum range. In addition, the following data were specified: the number of clangs to be generated in the computer run; the minimum and maximum durations of clangs (actual durations of successive clangs varied randomly within these limits); the number of voices to be generated in the clang; the probability of notes (vs. rests) occurring in each voice; and the range of frequency modulation for each voice. The instrument used is shown in Figure 9.

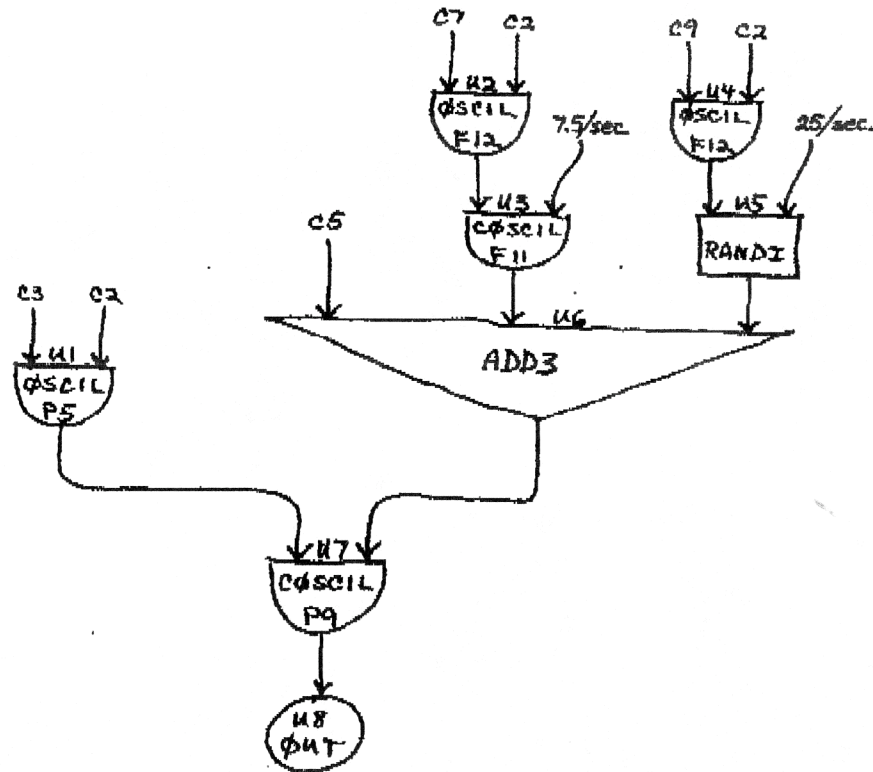


Figure 9. Instrument-design for use with PLF-2 (*Four Stochastic Studies*).

The program was run with various settings for clang-duration, number of voices and note-probability, and these tapes were later edited, becoming the *Four Stochastic Studies*. Much was learned from this first program, and each later program became more elaborate as it incorporated more refinements – greater flexibility, more precise controls, etc. However, these stochastic studies are remarkably interesting considering the simplicity of the program itself. I was well pleased with the results, while anxious to experiment with more elaborate compositional procedures.

One refinement, especially, seemed desirable. This was to make it possible to vary the large-scale mean-values in each parameter so that some sense of “direction” could be given to longer sequences, while still allowing the smaller details to vary randomly. In order to do this, and other things to be mentioned later, a new program (PLF 3) was written, whose input data included, for each section, initial and final values of the mean and two ranges in each parameter. The program first interpolates between the two values for the mean, according to the starting time of the clang in the section, then computes the clang-mean by adding to or subtracting from this mean a random number within the (first) specified range, and finally computes the successive note-values within the (second) range (around the clang-mean). The instruments used with PLF 3 were as diagrammed in Figure 10, and were designed to produce either tones or noise-bands. The probability of a sound’s being a noise (vs. a tone) is given among the input data. Three more parameters are variable in PLF 3, besides duration, amplitude and frequency. These are amplitude-modulation rate (which becomes noise bandwidth for faster rates), amplitude-envelope function-number, and wave-form function-number. The two types of stored functions are arranged in arbitrary “scales,” and controlled in essentially the same way the other parameters are. (The arrangement of the function-number scales is not entirely arbitrary: for wave-form, the spectra with more energy in the lower harmonics were given the lower scale-values, and for amplitude-envelope, those with the shorter rise-times were given the lower values. Thus, a sequence could change, gradually, from less to more “penetrating” and/or “percussive” timbres, for example.)

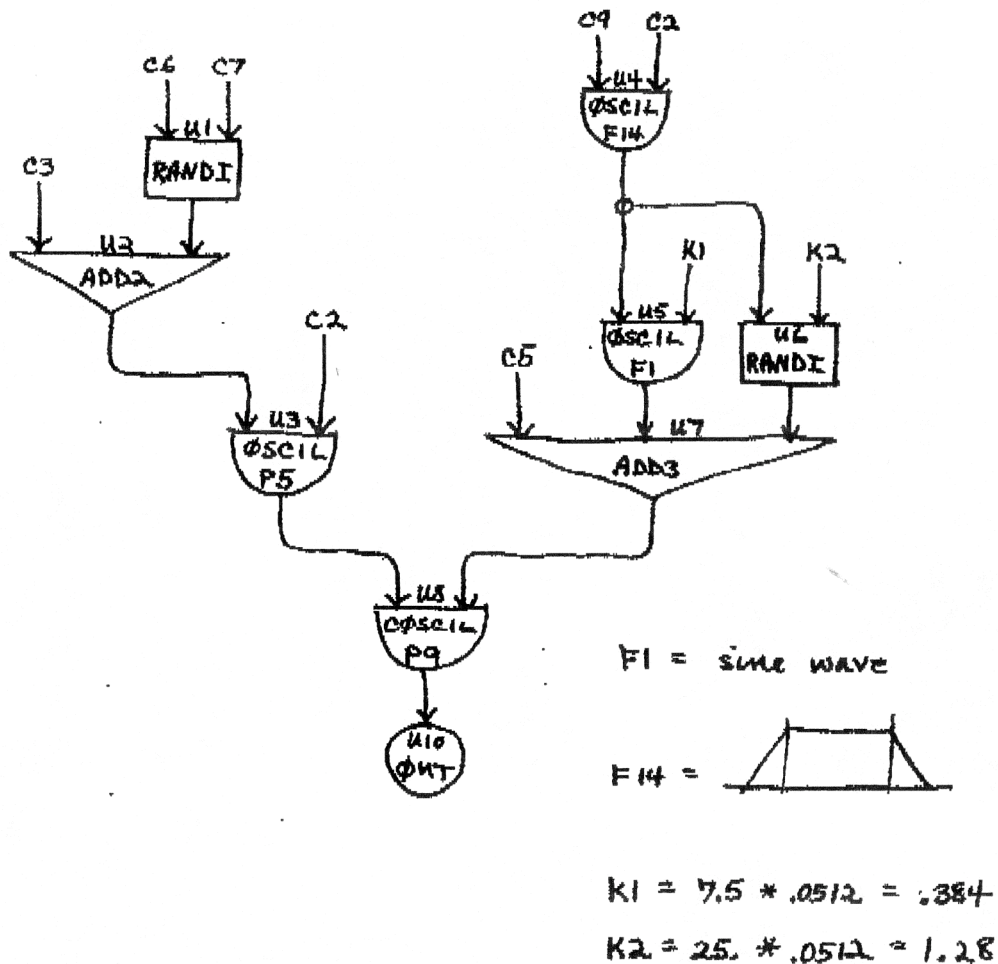


Figure 10. Instrument-design for use with PLF-3 (*Dialogue* and *Ergodos I*).

The PLF 3 subroutine was written in December, 1962, but the first composition (“Dialogue”) employing it was not completed until April, 1963, because another project was begun which had to be finished very quickly. This was the string quartet program, described in the following section (V). “Dialogue” was originally planned as a two-channel piece, with tones in one channel, noise-bands in the other. When the two tapes had been generated, however, I found the fixed correlation between

timbre and stereophonic position disturbing, so the two tapes were re-recorded into a single channel. The form of the piece is graphed in Figures 11a and 11b, which show the evolution of the large-scale mean values in each of the six parameters, as well as rest- and noise-probabilities and vertical density (number of voices generated per clang).

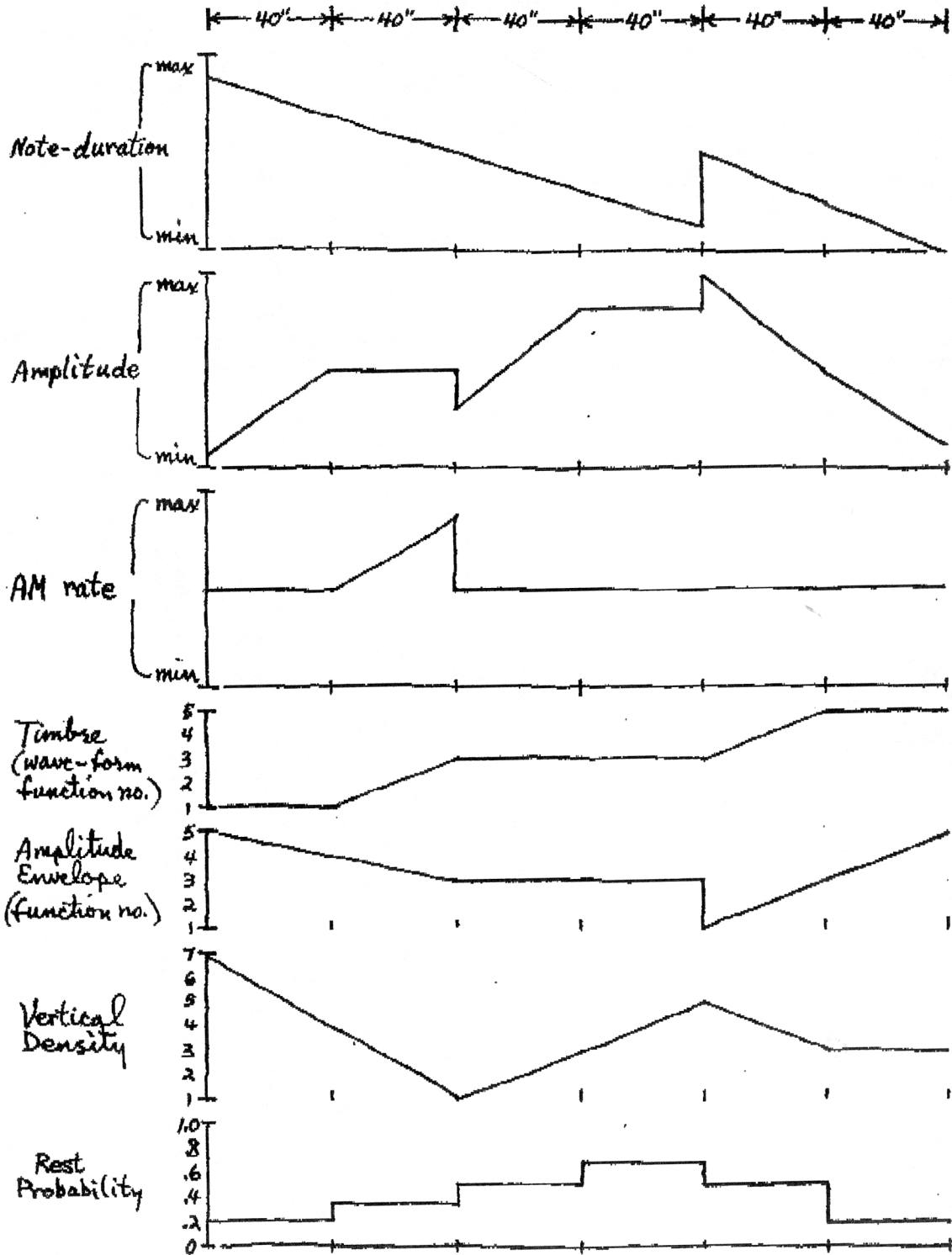


Figure 11a. Parametric Means for *Dialogue* (tonal stratum).

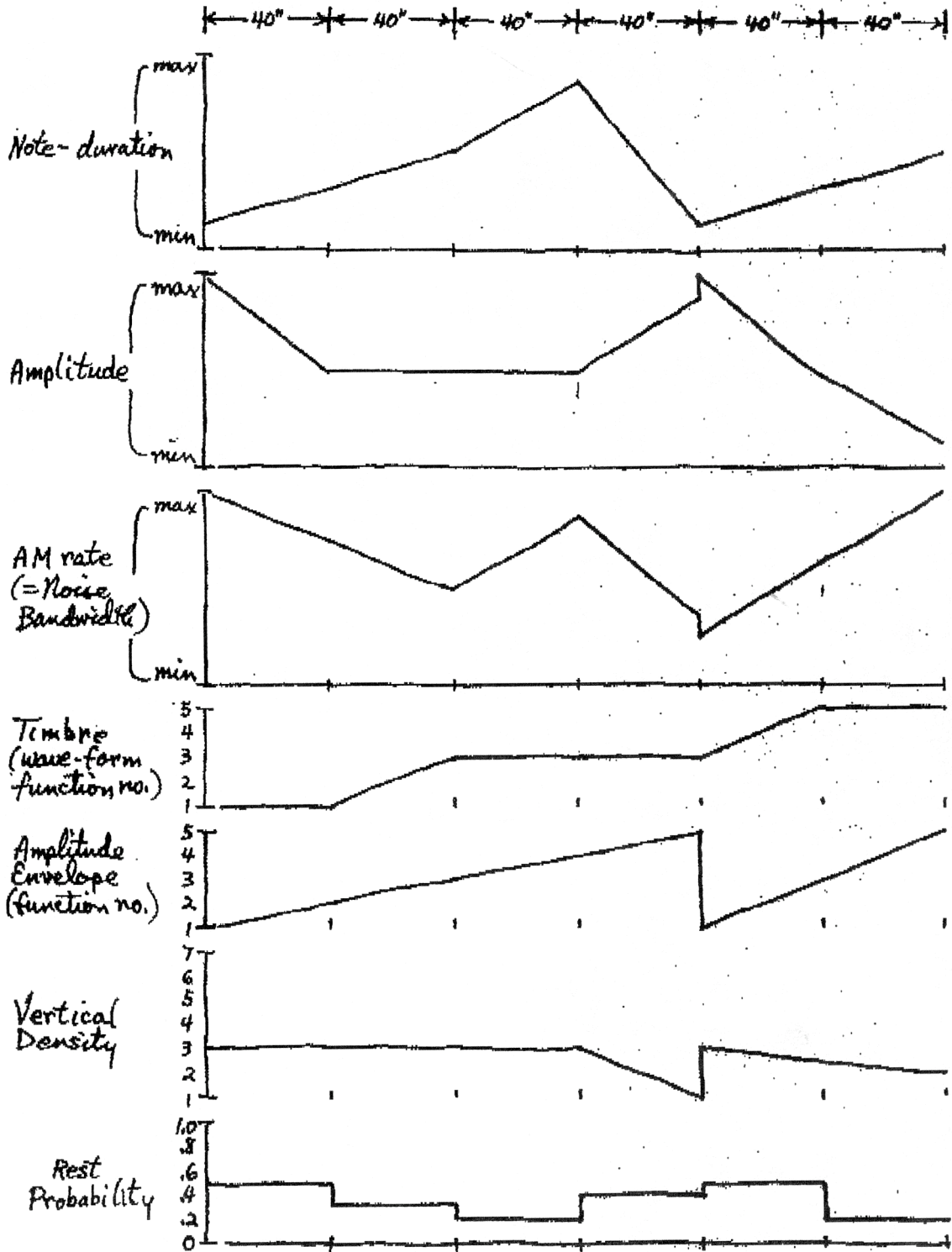


Figure 11b. Parametric Means for *Dialogue* (noise stratum).

### V. “Stochastic String Quartet”:

In December, 1962, there came a request for a computer-composed piece to be played by instruments (from the Paganini String Quartet in Los Angeles, who were to play the music on a special program celebrating “Science and Music” in February, 1963). Previously, such a use of the computer had only been attempted by Hiller at the University of Illinois (the “Illiac Suite”). One problem was involved that had not arisen in my earlier work with tape — how to quantize the various parameters of the sounds and print out the information in a way that could be transcribed into conventional musical notation. For most parameters, this problem was not great — pitches could be represented by integral numbers (of semi-tones, from the cello’s low C), dynamic levels by numbers from 1 (ppp) to 8 (fff), other parameters being encoded similarly. The real problem was time. With computer-generated sounds, I could deal with seconds and fractions of a second on a virtually continuous scale, with no necessary “rational” relationship between one note-duration and another. Conventional musical notation does not deal with time in this way, however, but rather in terms of measures which are integral multiples of a basic metrical unit duration, which may be subdivided, in turn, into various integral numbers of smaller units. In order to achieve as much variety as possible, within this system, I used the following procedure:

1. the duration of the metrical unit for the section was read from a card (giving the “tempo”);



2. the duration of each clang was computed as some integral multiple of this metrical unit duration (random within certain limits);
3. this clang-duration was next divided into some (limited random) number of “gruppetto” units, which may or may not equal the number of basic “metrical” units;
4. each of these secondary gruppetto units were further sub-divided into from one to three or four parts, yielding the (current) minimum possible note-value;
5. from the mean value and range of note-durations (computed along with corresponding values in other parameters for the clang as a whole, earlier in the program), a minimum and a maximum note-duration are computed;
6. for each note, the program steps through the smallest units, increasing the note-duration accumulatively, from the beginning to the end of the clang, testing the new duration after each addition; if the duration of the note is less than the minimum duration (described in (5) above), another increment is added to it, and it is tested again; if the duration is equal to or greater than the minimum, but less than the maximum duration for a note in that clang, the duration may be incremented or not (randomly, but with equal probability of either); if it is incremented, an indication that the note is “tied over” to the next unit is printed out; if it is not, the parameters for that note are printed out, and the program begins to compute a new note; finally, if the duration is

equal to or greater than the maximum duration, the note-parameters are of course printed out, as above. This process continues until all the subdivisions of each gruppetto unit, and all the gruppetto units themselves, for the clang, have been used up, for a given voice, and the next voice in the clang is computed.

The printout showed the number of metrical units in the clang, the number of gruppetto units and of the smaller unit in that gruppetto unit on which the note ended, and the transcription into musical notation was made using this information. Transcription turned out to be an exceedingly tedious process, however. In addition, the music was quite difficult to play (though no more difficult than some of Schoenberg's or Ives' music), and the Paganini Quartet ended up playing only a few pages of it. Later, the piece received a "reading" at the Bennington Composers' Conference, though the players refused to play the piece on the program it had been scheduled for. In the course of writing this program, another program was written that enabled the computer to read the "score" of the quartet and generate a tape version of the piece. The design of the computer "instruments" was done too quickly to make possible any very convincing simulation of the sounds of the ("real") stringed instruments, but the general rhythmic and textural character of the piece can be judged from this synthesized tape.

Since this first quartet was completed I have twice begun a new program for instrumental music, and twice abandoned the work before a piece was finished. The reasons for this were not clear to me until recently, and involve not only the experiences in writing the programs and listening to

the (synthetic) results on tape, but also the experiences in trying to get string-players to play the first quartet, and other, more general, changes in my musical attitudes in these last several months.

In the first quartet the complexities of the notated parts were such that a string player would have had to practice his part diligently, and even then the ensemble would probably have needed a conductor to keep it together. Now if every detail in the score were part of some “musical idea” (in a 19th century sense) that needed to be realized precisely, such a situation might be justified. But this was not the case. Each detail in the score was the result of a random selection process that was being used only to ensure variety, and might thus have been — within limits — anything else than what it was and still have fulfilled the conditions I had set up in the beginning. (At Bennington, I tried to explain this, and to assure the players that their “best approximation” to the part as notated was really sufficient. But the very appearance of the score itself contradicted me!). Thus, it began to be clear to me that there was an enormous disparity between ends and means in such a piece, and I have more recently tried to find a way to get that variety — in the “human,” instrumental situation — in ways more appropriate to the situation itself, in terms of the relationship between what the player sees and what he is expected to do.

Another problem arose with this quartet which has led to changes in my thinking and my ways of working, and may be of interest here. Since my earliest instrumental music (*Seeds*, in 1956), I have tended to avoid repetitions of the same pitch or any of its octaves before most of the other

pitches in the scale of 12 have been sounded. This practice derives not only from Schoenberg and Webern, and 12-tone or later serial methods, but may be seen in much of the important music of the century (Varèse, Ruggles, etc.). In the programs for both the *Quartet* and the *Dialogue*, steps were taken to avoid such pitch-repetitions, even though this took time, and was not always effective (involving a process of recalculation with a new random number, when such a repetition did occur, and this process could not continue indefinitely). In the quartet, a certain amount of editing was done, during transcription, to satisfy this objective when the computer had failed.

But several things about all this began to bother me: (a) it represented a kind of negative aspect of a process that was supposed to make “everything” possible; (b) it was a constraint applied only to one parameter — pitch — whereas almost all the other operations in the program were common to all parameters; and finally, (c) it used up a lot of computer-time (that might have been used to make more music, rather than less). Also, I had noticed that in the *Dialogue*, where the pitches are selected from a continuous scale (as opposed to the quantized scale of the *Quartet*), the pitch repetitions (two pitches within a very small interval of each other or of one’s octave) that got by the exclusion-process in the program did not seem to decrease the variability of the music, or interrupt the flow in the way they did in the *Quartet*. This suggested that the unison-octave avoidance was needed only when the pitch-scale was quantized as traditionally — only, that is, when the entropy of the pitch distribution had already been severely limited by such quantization. Accordingly, I no longer find it necessary to avoid any pitch, at the same

time that I intend never to leave undisturbed – even when working with instruments – the traditional quantized scale of available pitches. It is not too difficult to get around this with instruments (except for such as the piano) – it's mainly a matter of intention and resolve.

## VI. Ergodos I

Both the *String Quartet* and *Dialogue* made use of programming facilities enabling me to shape the large-scale form of a piece in terms of changing means and ranges in the various parameters in time. Now my thoughts took a different turn – an apparent reversal – as I began to consider what this process of “shaping” a piece really involved. Both the intention and the effect here were involved in one way or another with “drama” (as in Beethoven, say) – a kind of dramatic “development” that inevitably reflected (“expressed”) a guiding hand (mine), directing the course of things now here, now there, etc. What seemed of more interest than this was to give free reign to the sounds themselves, allowing anything to happen, within as broad a field of possibilities as could be set up. One question still remained as to the possible usefulness of my controls over the course of parametric means and ranges: are there ways in which the full extent and character of the “field” may be made more perceptible – more palpable – by careful adjustments of these values?

In later pieces, I was to test this question in various ways: by shaping only the beginning and the end of a piece, leaving the longer middle section “free” (*Ergodos I*), and by imposing a set of slowly oscillating functions on several parameters, with changing phase-relations between them in time (*Phases*). Finally, (in *Music for Player Piano* and *Ergodos II*), even these last vestiges of external “shaping” have disappeared, resulting in processes which evolve as freely as possible within the field of possibilities established for each one in the program itself. It is still often necessary to allow for a variable specification of parametric means and ranges (though

these no longer need to change in time), simply because it is still difficult to estimate the settings for these values that will result in the greatest variety and interest (while remaining within the practical limits imposed by the medium itself).

*Ergodos I* used the same composing program (PLF-3) and the same orchestra of computer instruments as *Dialogue*, but the nature of the music is very different. The composition consists of two, ten-minute monaural tapes that may be played either alone or together, either forward or backward. For each tape, only the first and last two minutes of the sound were subjected to any of the “shaping” of parametric means made possible by the composing program, and then only in a very simple way: the mean intensity begins (and ends) at a low level, and increases to mid-range toward the middle of the tape, while the mean tempo increases toward mid-range at one end of the tape (the beginning, say) increases away from the mid-range at the other (the end; if a tape is played in the reverse direction, the tempo decreases toward mid-range from the beginning, then decreases, further, away from mid-range at the end). During the middle six minutes of sound on each tape, all the parametric means are constant near the middle of their respective scale-ranges, and these ranges are at their maximum. Thus, the sounds on each tape are nearly ergodic, and thus the title — “*Ergodos*.”

In order to make possible so many different versions of this piece — so many alternative ways of performing it — it was necessary, first of all, to ensure a certain temporal symmetry, with respect to the amplitude envelope functions, for example. That is, first, there would have to be an

equal probability of envelope forms and their own “retrogrades.” And second, the average density of the sounds on each tape had to be great enough that a tape could be interesting when played by itself, and yet not so great that the two tapes could not be played together without losing clarity.

After preliminary tests to ascertain optimum settings of all parameters, and after generating the first two minutes of the first tape (the section with changing parameters) the program was run in one-minute segments. Each new segment on analog tape was then added to what had already been done, and I listened to the whole to determine whether more of these internal (constant) segments should be run before generating the final two minutes. My criterion was a subjective one that is not easy to define but was quite easily employed – does the “field of possibilities” seem to have been “used up”; does it seem that anything more can happen in this field that has not already happened? After the sixth of these constant, one-minute segments had been heard, it seemed to my ear that this criterion had been satisfied, and the final sections were generated.

For the second tape, the same number of sections were generated, so that both tapes would be of the same length. Before the second tape was begun, however, a few slight changes were made in certain parameters, adjustments that seemed needed after several hearings of the first. (My reactions were different when there was ten minutes of material than they had been in the testing period.) The final analog tapes were made by alternating between the sequence of digital tapes generated first, and the



second sequence, in order that the differences between the two series might be “balanced out” in the long run. Thus, the sounds on each tape are not truly ergodic, though my intention had been to make them as nearly so as possible (in the longer middle sections, at least), and they do approach this condition quite closely.

It may be of interest here to describe the changes that were made for the second set of digital tapes, as an example of the kind of values in various parameters that seem to approach the “mid-point” of the range, and of the extent of these ranges, but also to give an idea of the (small) magnitude of changes in statistical conditions that may have a perceptible musical effect. In the first set of digital tapes, the lower limit of the range of note-durations was  $1/16$  of a second, the upper limit 4 seconds. In the second set, this upper limit was increased to 5.3 seconds. In both cases, the overall mean-values were close to  $1/2$  second (log scales were used in nearly all parameters). In the first set of digital tapes, the note-rest probability (for the middle section) was .33, and four voices were generated per clang (average vertical density  $<3$ ). In the second set, this probability was increased to .5, and there were six voices per clang (ave. vertical density = 3, slightly greater than the density in the first set). Finally, the probability of a sound’s being a noise (rather than a tone) was .5 in the first set, .67 in the second. Settings in all other parameters were the same for the two series of digital tapes.

## VII. Phases and Ergodos II

In *Dialogue* and *Ergodos I* the variable parameters of the sounds were frequency, amplitude, AM rate (= noise bandwidth), wave-form and amplitude envelope form. The range of different timbres was thus relatively limited. In addition, each sound was either a tone or a noise-band, depending on the noise-probability specified for a sequence. In the next composing program, an attempt was made to extend the range of timbres as far as possible, and to achieve a continuous range of sound-qualities between these two extremes of tone and noise. I spent a great deal of time listening to all kinds of natural and mechanical sounds, as these occur in the environment, trying to determine their acoustical properties, and especially, the kinds of fluctuations in various parameters that were most often taking place within each sound. The whole “world” of environmental sounds (including sounds of musical instruments but no longer limited to these) became a kind of “model” for the range of sounds I wanted to be able to generate with the computer.

One of the most obvious aspects of many of these environmental sounds was their frequency instability – “glissandi” and “portamenti,” as well as faster modulations. The sounds in *Dialogue* and *Ergodos I* had some frequency modulation, but no frequency “enveloping,” and this now seemed a necessary extension of the list of variables. Filling in the gap between tones and noise-bands was achieved simply by allowing intermediate values to occur in the parameters affecting the noise – the range and rate of random amplitude modulation. In addition, it seemed

desirable to envelope the AM rate so that the bandwidth of the noise could vary within each sound.

In earlier orchestras, I had used a set of wave-form functions whose spectra contained formant peaks at different positions. The sounds of my “model” usually showed spectral variations independent of their fundamental frequency, which was not possible to achieve using such a fixed set of wave-form functions. What was clearly needed was the possibility of modifying the spectrum of each sound by means of a formant (band-pass) filter with continuously variable controls over center frequency and bandwidth, and the new instrument was designed accordingly. Since the current digital filter unit in the music compiler has a positive gain-factor greater than 1, varying as a function of both center frequency and bandwidth, it was necessary to compensate for this gain in the course of sample-generation. A FORTRAN function (RMSG) was written (based on computations made for me by Max Mathews and Jim Kaiser), which computes the rms gain of the filter (i.e., the ratio of the rms amplitude of the output to that of the input to the filter), and this function is “called” by the amplitude conversion functions (the CVT’s) used by the instrument. Figure 12 shows a block-diagram of the instrument, incorporating these changes; it is the instrument-design that was used for the piece called *Phases*.

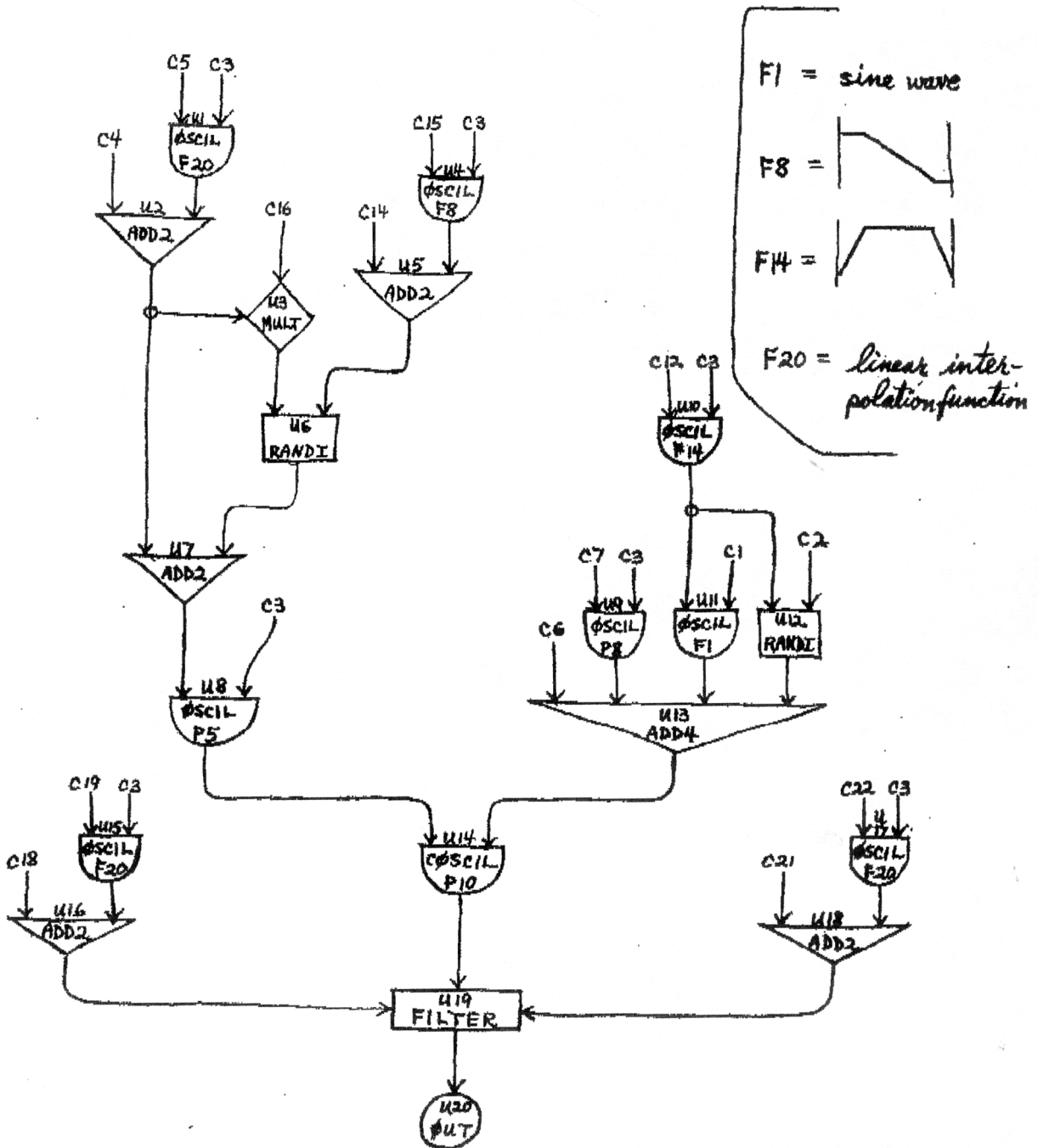


Figure 12. Instrument-design for *Phases*.

The composing program for *Phases* (PLF-5) also incorporated some new features. Whereas PLF-3 used random numbers to compute parametric values at two Gestalt levels (the means of each clang and of elements in a

clang), the parameters of each sequence (clang-group, the next larger Gestalt unit) could only be specified at the input. Since the input data usually referred to relatively long time-segments (30 to 90 seconds), Gestalt units of the order of the sequence (as perceived) were not actually being produced by the program. In the new program this was accounted for by including sequence-generation in the program in a way precisely analogous to the way clangs and elements were generated – via random numbers within a specified range above and below a larger mean value (in each parameter). The mean duration of clangs (and sequences) and a range of variability for these durations was specified in terms of a logarithmic time-scale (whereas in earlier programs, a minimum and maximum clang-duration had been specified, in terms of a linear time-scale). Parametric means and ranges were specified (for a section) using Mathews' CON function, so that fluctuations in these values could more easily be represented by straight-line segments than in earlier programs. Finally, no attempt was made to exclude unison or octave repetitions of pitch.

Figure 13 shows a graphic description of the most important variable parameters in *Phases*. The title “Phases” derives from the form of the piece, in which amplitude, note-duration and the noise-parameters were varied sinusoidally — oscillating around the mid-points of their respective scales at different rates, so that continually changing phase-relations between their mean values resulted in the course of the piece. By comparison with the ergodic form of *Ergodos I*, this was a small step backward — an experiment, really, to determine whether this kind of variation might produce a larger form more interesting than the ergodic

one, without sacrificing much in the way of variety. At this moment, the experiment remains inconclusive – I have not yet lived with these pieces long enough to be sure of my own reactions to them, in these large-formal terms.

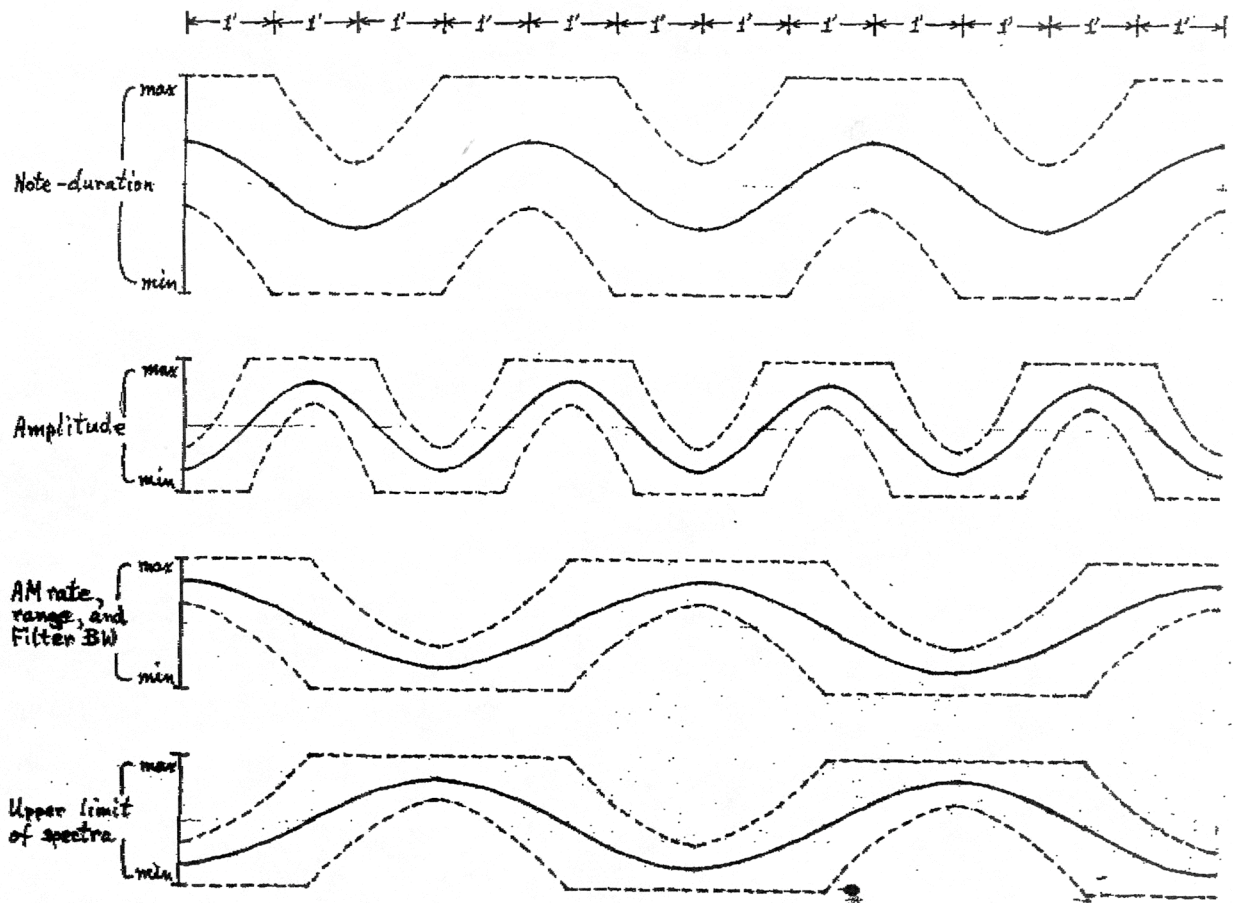


Figure 13. Parametric Means and ranges (dotted lines) for *Phases*.

*Phases* was completed in December 1963, and I began almost immediately to work on what was to become *Ergodos II*. Although provisions for stereophonic output have been incorporated in the Music Compiler since the summer of 1963, I had not yet made use of them. The need for stereophonic distribution of sounds had been apparent for a long time,

however, and I was determined to add this to the list of variables already active. Otherwise, the orchestra used for *Ergodos II* was almost identical to the one for *Phases*, with some minor revisions to improve the signal-to-noise ratio of the output (a problem caused by the digital filter). The form of the piece is ergodic again, without even the shaping of the beginning and end of the tape that was done in *Ergodos I*. The settings of the means and ranges of the various parameters were nearly the same as for *Ergodos I* and (the average means of) *Phases*, except that the rest-probabilities are higher — and there is thus a greater proportion of silence on the tape than in previous pieces. The final tape is eighteen minutes long, and may be played in either direction, beginning and ending at any points (i.e., a performance need not last the whole eighteen minutes). In addition, the tape might be subdivided into two or more segments of approximately equal length, and these segments played simultaneously (over one to N pairs of loudspeakers, for N segments). *Ergodos II* is the last composition to have been completed during my term at the Labs. Another piece was begun after its completion, but abandoned when my dissatisfaction with the early test results made it clear that I would not have time to complete it before leaving.